Published in final edited form as:

Brainlesion. 2018 ; 10670: 43–54.

Joint Intensity Fusion Image Synthesis Applied to Multiple Sclerosis Lesion Segmentation

Greg M. Fleishman¹, Alessandra Valcarcel², Dzung L. Pham³, Snehashis Roy³, Peter A. Calabresi⁴, Paul Yushkevich¹, Russell T. Shinohara², and Ipek Oguz¹

¹Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA

²Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA 19104, USA

³CNRM, The Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD 20817, USA

⁴Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

Abstract

We propose a new approach to Multiple Sclerosis lesion segmentation that utilizes synthesized images. A new method of image synthesis is considered: joint intensity fusion (JIF). JIF synthesizes an image from a library of deformably registered and intensity normalized atlases. Each location in the synthesized image is a weighted average of the registered atlases; atlas weights vary spatially. The weights are determined using the joint label fusion (JLF) framework. The primary methodological contribution is the application of JLF to MRI signal directly rather than labels. Synthesized images are then used as additional features in a lesion segmentation task using the OASIS classifier, a logistic regression model on intensities from multiple modalities. The addition of JIF synthesized images improved the Dice-Sorensen coefficient (relative to manually drawn gold standards) of lesion segmentations over the standard model segmentations by 0.0462 ± 0.0050 (mean $\pm$ standard deviation) at optimal threshold over all subjects and 10 separate training/testing folds.

1 Introduction

Multiple Sclerosis (MS) is a chronic, inflammatory, autoimmune disease of the central nervous system (CNS) wherein myelin sheaths which surround and insulate CNS nerves are damaged [7]. Regions of active inflammation and/or prior damage are visible in various magnetic resonance imaging (MRI) modalities as lesions. MS lesion segmentation has been an active area of research with many contributions including LesionTOADs [17], S3DL [15], IMaGe [19], and QuEEN [13,20], with many others recently reviewed [8] and evaluated [3]. However, no existing tool is a complete solution and improved lesion segmentation in MS

could lead to better characterization of disease trajectory and/or enable the observation of effective new therapies, either of which could improve patient outcomes.

Our aim is to improve MS lesion segmentation through image synthesis. Image synthesis is the process of artificially generating an image with geometry and appearance similar to those of a desired target [10]. Among the many approaches to image synthesis is the general idea of combining patches from real image acquisitions to synthesize an image with realistic contrast [16]. Image synthesis has been proposed as a means of assisting MS lesion segmentation, wherein for a given MS subject whose actual images contain lesions, an image is synthesized without lesions but with comparable features elsewhere [5].

Another method in which patches from different sources are fused to synthesize a final product is multi-atlas label fusion (MALF) [9]. In MALF, a library of atlases with associated segmentations are deformably registered to a target image. The segmentations for all atlases are resampled to the space of the target and fused according to a procedure that varies between methods to produce a final segmentation of the target. Another method which fuses patches is nonlocal means (NLM) filtering: non-local patches from within the same image are found and fused to reduce noise locally without blurring edges [1,2,22,23].

This work is inspired by the intersection of MALF and NLM, wherein corresponding patches from warped control atlases are fused rather than labels or patches from within the same image. These patches are local spatially, but nonlocal in that they derive from independent images. As a result, we synthesize images whose characteristics we can constrain by the choice of atlases. We refer to our proposed method as joint intensity fusion (JIF). The synthesized images are then utilized in MS lesion segmentation as additional features to a classifier. To summarize, our goal is to synthesize lesion-free images of MS subjects and test the hypothesis that the addition of such images to a lesion classifier would improve performance.

2 Methods

Our complete pipeline is summarized in Fig. 1. A library of healthy control atlases is registered and subsequently histogram matched to a target. The registered/histogram matched atlases and target are input to the Joint Label Fusion (JLF) algorithm, which produces a spatially varying weight map for each atlas. The atlases are combined according to these spatially varying weights to yield a synthesized image. Synthesized images are then treated as additional modalities to the OASIS multi-modality MS lesion classifier; we investigate whether their inclusion improves classification of lesions.

2.1 Data and Preprocessing

Our data set consists of 98 MS subjects (whose images we refer to as targets) and 29 healthy controls (whose images we refer to as atlases). Each subject has one scan from each of four modalities: T1-weighted (T1w), T2-weighted (T2w), proton density weighted (PD), and fluid attenuation inversion recovery (FLAIR). All scans were acquired on the same 3T MRI scanner (a Phillips scanner at the Johns Hopkins Hospital). Each MS subject also has a gold standard lesion mask segmented manually by a neuroscientific researcher.

For all subjects, all images were rigidly aligned to the Montreal Neurological Institute (MNI) standard space (voxel resolution 1 mm³) [12]. We also applied the N3 inhomogeneity correction algorithm [18] to all images and removed extracerebral voxels using SPECTRE, a skull-stripping procedure [4].

2.2 Registration and Intensity Normalization

For each modality in parallel, all preprocessed atlases were registered to each preprocessed target. Registration included three steps: a 6 parameter rigid alignment, a 12 parameter affine alignment, and a deformable registration. Registrations were done with the package *greedy* which is described in Yushkevich et al. [25]. We normalize the non-standardized image intensity histograms by histogram matching each registered atlas to its corresponding target.

2.3 Joint Intensity Fusion

Our image synthesis approach is inspired by the multi-atlas label fusion (MALF) method for image segmentation. In MALF, a target image is segmented by registering to it a library of segmented atlases. All atlas labels are resampled to the space of the target; producing a sampling of labels at each voxel. This sampling is fused into a consensus segmentation with fusion strategies that often involve spatially varying weights based on image similarity metrics [9]. The primary methodological contribution of this work is to use a MALF-like strategy to fuse MRI signal directly rather than labels.

In particular, we chose to apply the joint label fusion (JLF) MALF strategy which has proven a successful means to achieve MALF segmentation, rated the best of nearly 25 methods in the 2012 MICCAI Grand Challenge and Workshop on Multi-Atlas Labeling [11]. JLF fuses labels by weighted averaging, the distribution of weights over atlases varies with position. However the definition of the weights accounts for correlations between atlases in the library [24]. JLF chooses weights that minimize the expected squared difference between the (unknown) ground truth label and the estimated label. To do so, JLF requires a matrix:

$$M_x(i, j) = p(\delta^i(x)\delta^j(x) = 1 \mid I_T, A_1, \dots, A_n) \quad (1)$$

whose ij^{th} entry is the probability that both atlases A_i and A_j mislabel location x in the target image I_T , δ

$\delta(x) \in \{0, 1\}$ is the label error, 0 if $A_i(x)$ gives the correct label and 1 otherwise. M_x captures the correlations between atlases in the library. However, M_x is unknown and must be approximated. Let $\mathcal{N}(x)$ be a neighborhood around voxel x ; JLF approximates this matrix as:

$$p(\delta^i(x)\delta^j(x) = 1 \mid I_T(y), A_i(y), A_j(y) \mid y \in \mathcal{N}(x)) \quad (2)$$

$$= \left[\sum_{y \in \mathcal{N}(x)} \mid I_T(y) - A_i(y) \mid \mid I_T(y) - A_j(y) \mid \right]^\beta$$

We apply the JLF technique to the set of registered healthy control atlases and MS targets, but use the resultant weights to fuse the atlases themselves rather than labels. That is, we synthesize an image where each location is the weighted average of the warped healthy control atlases, with the weights given by JLF. Recall, JLF weights vary spatially. We call this method of image synthesis Joint Intensity Fusion (JIF). Examples of JIF synthesized images are shown in the second column of Fig. 4.

2.4 MS Lesion Classification and Evaluation

We trained two versions of OASIS MS lesion classifiers [21]: one using only the four target MS subject modalities (FLAIR, T1w, T2w, and PD) and another using those same modalities and their JIF synthesized versions. OASIS is a logistic regression model for the probability that any voxel V is in a lesion ($P(V = \text{lesion})$). The logit function of the model is:

$$\begin{aligned} \text{logit}(P(V = \text{lesion})) = & \beta_0 + \beta_1 FLAIR + \beta_2 T1 + \beta_3 T2 + \beta_4 PD \\ & + \beta_5 FLAIR \times FLAIR_{\sigma_1} + \beta_6 T1 \times T1_{\sigma_1} + \beta_7 T2 \times T2_{\sigma_1} + \beta_8 PD \times PD_{\sigma_1} \\ & + \beta_9 FLAIR \times FLAIR_{\sigma_2} + \beta_{10} T1 \times T1_{\sigma_2} + \beta_{11} T2 \times T2_{\sigma_2} + \beta_{12} PD \times PD_{\sigma_2} \end{aligned} \quad (3)$$

where each modality is represented by its name, X_{σ_1} and X_{σ_2} are versions of modality X smoothed with isotropic Gaussians of standard deviations σ_1 and σ_2 respectively, and the β_i are regression coefficients. We augment this function to include the JIF synthesized images as if they were four new modalities:

$$\begin{aligned} \text{logit}(P(V = \text{lesion})) = & \beta_0 + \beta_1 FLAIR + \beta_2 T1 + \beta_3 T2 + \beta_4 PD \\ & + \beta_5 \mathbf{FLAIR^s} + \beta_6 \mathbf{T1^s} + \beta_7 \mathbf{T2^s} + \beta_8 \mathbf{PD^s} \\ & + \beta_9 FLAIR \times FLAIR_{\sigma_1} + \beta_{10} T1 \times T1_{\sigma_1} + \beta_{11} T2 \times T2_{\sigma_1} + \beta_{12} PD \times PD_{\sigma_1} \\ & + \beta_{13} FLAIR \times FLAIR_{\sigma_2} + \beta_{14} T1 \times T1_{\sigma_2} + \beta_{15} T2 \times T2_{\sigma_2} + \beta_{16} PD \times PD_{\sigma_2} \\ & + \beta_{17} \mathbf{FLAIR^s} \times \mathbf{FLAIR^s}_{\sigma_1} + \beta_{18} \mathbf{T1^s} \times \mathbf{T1^s}_{\sigma_1} + \beta_{19} \mathbf{T2^s} \times \mathbf{T2^s}_{\sigma_1} + \beta_{20} \mathbf{PD^s} \times \mathbf{PD^s}_{\sigma_1} \\ & + \beta_{21} \mathbf{FLAIR^s} \times \mathbf{FLAIR^s}_{\sigma_2} + \beta_{22} \mathbf{T1^s} \times \mathbf{T1^s}_{\sigma_2} + \beta_{23} \mathbf{T2^s} \times \mathbf{T2^s}_{\sigma_2} + \beta_{24} \mathbf{PD^s} \times \mathbf{PD^s}_{\sigma_2} \end{aligned} \quad (4)$$

where X^s is the synthesized version of modality X ; new terms are bolded. We refer to the control model (3) as OASIS(raw) and the experimental model (4) as OASIS(raw+synth).

2.5 Experiments and Evaluation

We compare OASIS(raw) and OASIS(raw+synth) with ten-fold cross validation. Each fold consists of 20 randomly selected training subjects. The remaining 78 subjects are tested. For

each test subject, OASIS provides a probability map (p-map), i.e. each voxel is assigned a probability $p \in [0, 1]$ where $p = 0$ implies certainty of non-lesion and $p = 1$ implies certainty of lesion. We threshold the p-maps at all values in $\{0, 0.01, \dots, 1.0\}$ and report the Dice-Sorensen coefficient (DSC) and the receiver operating characteristic (ROC) curve to evaluate performance. The significance of the difference between OASIS(raw+synth) and OASIS(raw) curves from corresponding cross validation folds was assessed with the DeLong test [6]. A balanced and random sub-selection of true positive and true negative voxels were selected from each subject to satisfy consistency with the DeLong test assumption of independent samples.

Additionally, we include results on the same 10 testing folds computed with the LesionTOADS method [17]. LesionTOADS is an unsupervised algorithm to segment brain tissues and lesions simultaneously. Unlike OASIS which requires training data for lesion segmentation, LesionTOADS uses fuzzy c-means to segment the T1w and FLAIR image intensities into multiple tissue classes such as grey matter, white matter, cerebrospinal fluid, and lesions. Further, Lesion-TOADS enforces topology checks so that the resultant segmentation is topologically similar to a template. LesionTOADS produces fuzzy memberships of lesions, which are thresholded at $\{0, 0.01, \dots, 1.00\}$ to produce binary lesion segmentations. Although LesionTOADS can be tuned with different parameters [14] for different expected lesion load present in a subject, we processed all subjects with the default parameter set for uniform comparison with OASIS.

Because they are averages of multiple acquisitions, the synthetic images have higher signal to noise ratio (SNR) than the raw acquisitions. Thus, it is necessary to investigate whether any differences in lesion segmentation are solely due to the higher SNR, or some other property of the synthetic images. We test this by using Unbiased Non-Local Means (UNLM) [22], which averages non-local neighborhoods that are similar in appearance to the current patch, as an alternative smoothing method. For every target MS image, we generated a smoothed version with UNLM. We then trained OASIS models using the target images and their UNLM smoothed counterparts; we refer to these models as OASIS(raw+unlm). We compare the performance of OASIS(raw+synth) and OASIS(raw+unlm).

3 Results

The ROC and partial ROC (pROC) curves (up to false positive rates of 0.01) are shown in Fig. 2. False positive rates above 0.01 are not clinically useful for MS lesion segmentation; consider, 1% of the volume of a healthy control subject is on the order of 10 cm^3 , which is commensurate with the average lesion load of an MS subject. The ROC and pROC curves show that OASIS(raw+synth) has higher true positive rates for any given false positive rate including those in the clinically relevant range. The DeLong test showed that the improvement in the OASIS(raw+synth) model was significant with $p < 2.2 \times 10^{-16}$ in all 10 folds.

DSC as a function of threshold is shown in Fig. 3. The differences between corresponding OASIS(raw+synth) folds and OASIS(raw) folds are non-negative, indicating that including the JIF synthesized images results in higher DSC for any threshold. Table 1 shows the mean

and median (over subjects) DSC obtained at optimal threshold for each fold by OASIS(raw+synth), OASIS(raw), and the LesionTOADs package [17]. For each fold, OASIS(raw+synth) obtains the highest mean and median DSC. Figure 4 shows 3 qualitative results illustrating the behavior of the two models.

Figure 5 contrasts OASIS(raw), OASIS(raw+unlm), and OASIS(raw+synth) and is analogous to Figs. 2 and 3. It shows that the addition of UNLM smoothed images to OASIS is not sufficient to recover the improvements gained by the OASIS(raw+synth) model, i.e. the improvements gained in OASIS(raw+synth) are not likely due solely to the increased SNR in the synthetic images.

4 Discussion and Conclusions

The results indicate that including the JIF synthesized images in the OASIS model improves lesion segmentation relative to the standard OASIS model. Inspection of Eq. (4) illuminates the improvement. Only features present in the library of atlases can be reconstructed in a synthesized image. As the atlas library is composed of only healthy controls, it is unlikely to synthesize a region with appearance similar to a lesion. Equation (4) contains linear combinations of the MS target images and their synthesized counterparts. Hence, what is present in the target images but missing in the synthesized images can be accentuated.

We investigated which of the input features were most useful to the model. The magnitude of the coefficients β_i in Eq. (4) are an indicator of the importance of their corresponding feature. However, these coefficients must be normalized to the scale of their features. Figure 6 shows boxplots (over the 10 cross-validation folds) of the coefficients of the OASIS(raw+synth) model, where each coefficient is multiplied by the median value of its corresponding feature. Those coefficients furthest from zero (either close to the top or bottom of the plot) were the most discriminatory between lesion and non-lesion voxels. Noticeably, many of the interaction terms are in the center, and are the weakest contributors. FLAIR and T1w modalities, both raw and synthetic, are the most discriminatory.

The proposed method involves multiple registrations, histogram matching, JIF computation of weights, fusion, and subsequent training and testing of a classifier model; we report approximate runtimes for those steps: **registration**: 300s per image pair, **HM**: 5s per image pair, **JIF**: 540s total for 29 atlases, **fusion**: 10s per synthetic image, and **OASIS**: 60s per image.

The registration package *greedy* [25] is, compared to other publicly available packages, a very fast implementation of diffeomorphic image registration. The time reported above represents rigid, affine, and deformable registrations done in sequence. As with any gradient descent based optimization, run time will depend on the quality of the initialization, i.e., how well matched the input images are initially. For JIF, the run time is a function of the number of atlases; the reported time was computed using 29 atlases.

OASIS training fits a logistic regression model and OASIS testing evaluates that model; the time for both computations is constant given a fixed number of inputs. However, OASIS does feature selection prior to training and testing, the most computationally expensive

portion of which is smoothing the input images multiple times. We thus only report the time required to perform feature selection for one image. When training using N input modalities and M subjects, the total training time is approximated by $N \times M \times R$ where R is the number reported above.

In closing, inclusion of JIF synthetic images significantly improves lesion classification of the OASIS mode and with the inclusion of synthetic images, OASIS outperformed other state-of-the-art methods. Future work will investigate alternative classifier models and additional applications for JIF synthesized images.

Acknowledgments

This work supported by the National Institutes of Health grants: R01-NS094456, R01-EB017255, R01-NS085211, R21-NS093349, R01-NS082347, R01-NS070906 and by the National Multiple Sclerosis Society grant: RG-1507-05243.

References

1. Awate S, Whitaker R. Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Trans PAMI*. 28:364–376.2006;
2. Buades A, Coll B, Morel JM. Nonlocal image and movie denoising. *Int J Comput Vis*. 76(2):123–139.2008;
3. Carass A, et al. Longitudinal multiple sclerosis lesion segmentation: resource and challenge. *NeuroImage*. 148:77–102.2017; [PubMed: 28087490]
4. Carass A, Cuzzocreo J, Wheeler MB, Bazin PL, Resnick SM, Prince JL. Simple paradigm for extra-cerebral tissue removal: algorithm and analysis. *NeuroImage*. 56(4):1982–1992.2011; [PubMed: 21458576]
5. Cardoso MJ, Sudre CH, Modat M, Ourselin S. Template-based multimodal joint generative model of brain data. *IPMI*. :17–29.2015
6. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 44:837–845.1988; [PubMed: 3203132]
7. Dendrou CA, Fugger L, Friese MA. Immunopathology of multiple sclerosis. *Nat Rev Immunol*. 15(9):545–558.2015; [PubMed: 26250739]
8. Garcia-Lorenzo D, Francis S, Narayanan S, Arnold DL, Collins DL. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Med Image Anal*. 17(1):1–18.2013; [PubMed: 23084503]
9. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: a survey. *Med Image Anal*. 24(1):205–219.2015; [PubMed: 26201875]
10. Jog A, Roy S, Carass A, Prince JL. Magnetic resonance image synthesis through patch regression. *Proceedings of IEEE ISBI 2013*. :350–353.2013
11. Landman BA, Warfield SK. MICCAI 2012 workshop on multi-atlas labeling. *MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling Challenge Results*. 2012
12. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A probabilistic atlas of the human brain: theory and rationale for its development. *NeuroImage*. 2(2):89–101.1995; [PubMed: 9343592]
13. Mejia AF, Sweeney EM, Dewey B, Nair G, Sati P, Shea C, Reich DS, Shinohara RT. Statistical estimation of T1 relaxation times using conventional magnetic resonance imaging. *NeuroImage*. 133:176–188.2016; [PubMed: 26732403]
14. Roy S, Agarwal H, Carass A, Bai Y, Pham DL, Prince JL. Fuzzy c-means with variable compactness. *IEEE International Symposium on Biomedical Imaging*. 2008

15. Roy, S, Carass, A, Prince, JL, Pham, DL. Subject specific sparse dictionary learning for atlas based brain MRI segmentation. In: Wu, G, Zhang, D, Zhou, L, editors. MLMI 2014. LNCS. Vol. 8679. Springer; Cham: 2014. 248–255.
16. Tsafaris, SA, Gooya, A, Frangi, AF, Prince, JL, editors. SASHIMI 2016. LNCS. Vol. 9968. Springer; Cham: 2016.
17. Shiee N, Bazin PL, Ozturk A, Reich DS, Calabresi PA, Pham DL. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *NeuroImage*. 49(2):1524–1535.2010; [PubMed: 19766196]
18. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 17(1):87–97.1998; [PubMed: 9617910]
19. Subbanna N, Precup D, Arnold D, Arbel T. Image: iterative multilevel probabilistic graphical model for detection and segmentation of multiple sclerosis lesions in brain MRI. *IPMI*. :514–526.2015
20. Suttner L, Mejia A, Dewey B, Sati P, Reich D, Shinohara R. Statistical estimation of white matter microstructure from conventional MRI. *NeuroImage: Clinical*. 12:615–623.2016; [PubMed: 27722085]
21. Sweeney EM, Shinohara RT, Shiee N, Mateen FJ, Chudgar AA, Cuzzocreo JL, Calabresi PA, Pham DL, Reich DS, Crainiceanu CM. OASIS is automated statistical inference for segmentation, with applications to multiple sclerosis lesion segmentation in MRI. *NeuroImage: Clinical*. 2:402–413.2013; [PubMed: 24179794]
22. Tristán-Vega A, García-Pérez V, Aja-Fernández S, Westin CF. Efficient and robust nonlocal means denoising of MR data based on salient features matching. *Comput Methods Programs Biomed*. 105(2):131–144.2011; [PubMed: 21906832]
23. Tustison N, Avants B, Wang H, Xie L, Coupe P, Yushkevich P, Manjon J. A patch-based framework for new ITK functionality: Joint fusion, denoising, and non-local super-resolution. *Insight Journal*. 2017
24. Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. Multi-atlas segmentation with joint label fusion. *IEEE Trans PAMI*. 35(3):611–623.2013;
25. Yushkevich PA, et al. Fast automatic segmentation of hippocampal subfields and medial temporal lobe subregions in 3 Tesla and 7 Tesla T2-weighted MRI. *Alzheimer's & Dement J Alzheimer's Assoc*. 12(7):P126–P127.2016;

Fig. 1. Joint Intensity Fusion Image Synthesis for MS lesion segmentation. Manual segmentations were used when training the OASIS models; for test subjects, they were used for evaluation only.

Fig. 2.

ROC and pROC curves for OASIS(raw+synth), OASIS(raw), and Lesion-TOADs classifier models. All 10 folds and their average are depicted. The OASIS(raw+synth) curves dominate the OASIS(raw) curves. The LesionTOADs curve only exceeds OASIS(raw) and OASIS(raw+synth) for very high false positive rates that have limited value in practice. The inset shows that OASIS(raw+synth) performance exceeds both alternatives within the clinically relevant range.

Fig. 3.

DSC over probability map threshold for OASIS(raw), OASIS(raw+synth), and LesionTOADs classifier models. All 10 folds and their average are depicted. The difference curves are OASIS(raw+synth) - OASIS(raw) and are non-negative, indicating the OASIS(raw+synth) curves are everywhere greater than their OASIS(raw) counterparts. Both OASIS(raw+synth) and OASIS(raw) have higher peak DSC than LesionTOADs.

Fig. 4.

Acquired FLAIRs with manual segmentation boundary, corresponding synthesized FLAIRs, and manual segmentation masks overlaid with lesion probability maps for OASIS(raw+synth) and OASIS(raw). **First row:** true positive where OASIS(raw+synth) gives high probability, but OASIS(raw) does not. **Second row:** true positive where OASIS(raw+synth) has higher probability than OASIS(raw). **Third row:** true negative where OASIS(raw+synth) has lower probability than OASIS(raw).

Fig. 5.

UNLM: Unbiased Non-Local Means smoothing. For ROC, pROC, and DSC, OASIS(raw+unlm) appears to perform only slightly better than OASIS(raw) and visibly worse than OASIS(raw+synth); see Table 1 also.

Fig. 6.

Coefficients β_i from Eq. (4) multiplied by the median of their corresponding feature; boxplot is over all 10 training folds. If the median is between ± 0.2 then the box is red; blue and green depict positive and negative contributors respectively. Noticeably, the interaction features tend to be the least predictive, whereas FLAIR and T1w features (both raw and synthetic) tend to be the most predictive. (Color figure online)

Table 1

Mean and median DSC for all models at peak threshold for each fold.

MEAN of Fold:	1	2	3	4	5	6	7	8	9	10	Mean (Std. dev.)
OASIS(raw+synth)	0.56	0.58	0.59	0.56	0.56	0.57	0.57	0.55	0.59	0.57	0.57 (0.01)
OASIS(raw)	0.53	0.53	0.54	0.51	0.52	0.53	0.53	0.50	0.53	0.52	0.52 (0.01)
LesionTOADs	0.42	0.44	0.44	0.43	0.42	0.43	0.42	0.42	0.44	0.42	0.43 (0.01)
OASIS(raw+unlm)	0.54	0.54	0.55	0.53	0.53	0.54	0.54	0.52	0.54	0.53	0.53 (0.01)
MEDIAN of Fold:	1	2	3	4	5	6	7	8	9	10	Median (IQR)
OASIS(raw+synth)	0.64	0.63	0.63	0.64	0.62	0.64	0.63	0.61	0.64	0.64	0.64 (0.63, 0.64)
OASIS(raw)	0.60	0.58	0.58	0.58	0.59	0.59	0.60	0.57	0.59	0.60	0.59 (0.58, 0.60)
LesionTOADs	0.46	0.48	0.50	0.50	0.48	0.50	0.46	0.48	0.50	0.48	0.48 (0.48, 0.50)
OASIS(raw+unlm)	0.61	0.58	0.59	0.60	0.59	0.61	0.62	0.61	0.61	0.62	0.61 (0.60, 0.61)